计算机类本科规划教材

计算方法

（第2版）

◆ 李桂成 编著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

计算机类本科规划教材

计 算 方 法

（第2版）

李桂成　编著

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书比较全面地介绍了现代科学与工程计算中常用的数值计算方法。全书共分 12 章，主要内容有：引论、计算方法的数学基础、方程求根、解线性方程组的直接法、解线性方程组的迭代法、函数插值、函数逼近、数值积分与数值微分、常微分方程初值问题的数值解法、矩阵特征值计算、函数优化计算和 MATLAB 编程基础及其在计算方法中的应用。

本书知识体系完整，从简要回顾与计算方法有关的数学基础知识，到介绍现代计算软件 MATLAB 在本领域中的应用，书中每个算法都配有结构化流程图，大部分算法给出了 MATLAB 语言和 C 语言的源代码，书后附有上机实验题目。可从华信教育资源网（www.hxedu.com.cn）免费下载的教学资源包括：电子教案、各章习题解答和模拟试题。

本书可作为高等院校理工科计算机、电子信息类及近电类本科和研究生教材使用，也可供从事科学与工程计算的科技工作者和研究人员参考。

图书在版编目（CIP）数据

计算方法 / 李桂成编著. —2 版. —北京：电子工业出版社，2013.8
计算机类本科规划教材
ISBN 978-7-121-20328-2

Ⅰ. ①计… Ⅱ. ①李… Ⅲ. ①计算方法—高等学校—教材 Ⅳ. ①0241

中国版本图书馆 CIP 数据核字（2013）第 094733 号

责任编辑：冉 哲
印　　刷：北京市海淀区四季青印刷厂
装　　订：北京市海淀区四季青印刷厂
出版发行：电子工业出版社
　　　　　北京市海淀区万寿路 173 信箱　邮编　100036
开　　本：787×1 092　1/16　印张：20.25　字数：518.4 千字
印　　次：2013 年 8 月第 1 次印刷
印　　数：3 000 册　　定价：41.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

前　言

随着科学技术的飞速发展和计算机的广泛应用，现代科学已呈现出理论科学、实验科学和计算科学三足鼎立的局面。作为计算科学的重要手段和工具，计算方法已成为当代理工科大学生必备的基础和技能。从 20 世纪 80 年代起，“计算方法”就成为信息和计算机等专业本科生的专业基础课。作者从事“计算方法”教学 20 余载，在教学过程中发现，学生在学习“计算方法”课程时，存在三个明显的问题：一是学生以前学过的数学知识大多已淡忘，翻开《高等数学》和《线性代数》课本，也不知从何看起；二是在实践环节中，学生用以前学过的程序设计语言编写计算方法实验程序，感到非常吃力；三是许多学生对计算方法中的算法，只知其然，不知其所以然。鉴于此，作者将自己一直使用的讲义进行整理，编写了这本教材，以尝试解决这些问题。

本书自 2005 年出版以来，深受读者喜爱和同行专家关注，作者也在本科和研究生教学中使用本书。经过多年的教学实践，收到了比较满意的效果，总体反映良好，但也发现一些有待改进之处。为了更好地适应教学和自学的需要，作者认真听取了关心本书的同行专家和读者的建议，决定修订再版。

本书第 2 版的内容主要有以下变动：

（1）为了适应不同专业和层次的教学需求，增加了函数逼近、矩阵特征值计算和函数优化计算三章内容，在教学中可根据实际情况灵活调整。

（2）删除了解线性方程组的平方根法和超松弛迭代法、函数插值的低阶含导数项的插值、解常微分方程的线性多步法等内容。

（3）本着通俗易懂、利于教学、方便使用的原则，对部分内容进行了改写或重新编排。

（4）对附录 A 中的计算方法实验进行了补充、细化和完善。

（5）增加或改写了一些算法、程序、测试例题和部分习题。

全书共分 12 章：

第 1 章引论介绍计算方法与经典数学的差异，以及误差理论基础，主要包括误差的度量、分类、传播和数值计算的若干原则；

第 2 章简明、系统地介绍计算方法的数学基础，主要包括微积分、微分方程和线性代数的有关概念和结论；

第 3 章介绍非线性方程的数值解法，主要包括二分法、迭代法和牛顿法；

第 4 章介绍解线性方程组的直接法，主要包括消去法、矩阵分解法和误差估计；

第 5 章介绍解线性方程组的迭代法，主要包括雅可比迭代法、高斯-塞德尔迭代法，以及迭代公式的收敛性；

第 6 章介绍插值函数，主要包括拉格朗日插值、牛顿插值、埃尔米特插值和样条插值；

第 7 章介绍函数逼近，主要包括最佳一致逼近、最佳平方逼近及离散数据的曲线拟合；

第 8 章介绍数值积分和数值微分，主要包括牛顿-柯特斯求积公式、高斯求积公式、龙贝格求积公式及插值型求导公式；

第 9 章介绍常微分方程初值问题的数值解法，主要包括欧拉方法及其变形、龙格-库塔方法及单步法的收敛性和稳定性；

第 10 章介绍矩阵特征值计算，主要包括幂法、QR 方法及雅克比方法；

第 11 章介绍函数优化计算，主要包括一元函数优化计算、多元函数优化计算；

第 12 章介绍 MATLAB 编程基础及其在计算方法中的应用，主要包括 MATLAB 编程环境、矩阵计算、图形功能，以及用 MATLAB 实现计算方法中的基本算法，并对这些算法配有 MATLAB 和 C 语言源代码。

讲授全书内容，需要 40～60 学时，另外，实验需要 12～18 学时。

本书的特点是：

（1）知识结构完整。既有计算方法所需的数学知识，又有现代数值计算软件 MATLAB 在计算方法中的应用，一书在手，无须其他资料。

（2）注重实践和应用。书中提供约 40 个 MATLAB 和 6 个 C 语言源代码实例，以实现本书介绍的主要算法，并对结果进行比较，测试数据全部选自本书的例题；附录 A 中有精心设计的 11 个实验，供读者选用。

（3）方便教学。全书每章都有内容提要、教学建议和本章小结，说明每章的重点、难点、选学内容和所需教学时数。书后附录 A 中的 11 个实验，可以加强计算方法课程的实践环节。本书还配有电子教案、模拟试题和各章习题解答，供教师使用，可以登录华信教育资源网（www.hxedu.com.cn）注册后免费下载。

（4）便于自学。本书注重介绍算法的来龙去脉、基本思路和推导过程，每个算法都配有结构化流程图，并在第 12 章中有示例源代码和测试例题供读者参考。

特别值得一提的是，书中对比了 C 语言和 MATLAB 实现各算法的优劣，通过这种对比，读者有望深刻理解“计算方法”的精髓。

本书可作为大学本科计算机、电子信息和近电类专业的教材使用，也可供从事科学与工程计算的科技工作者和研究人员参考。

在本书第 2 版出版之际，作者特别感谢梁吉业教授对本书的长期关注和大力支持，感谢他对本书提出的宝贵意见和许多合理化的建议。电子工业出版社的高级编辑冉哲对本书的出版做了大量工作，在此表示感谢。

由于作者水平有限，书中难免有错误和疏漏之处，恳请读者指正。

作　者

反侵权盗版声明

举报电话：（010）88254396；（010）88258888

传　　真：（010）88254397

E-mail：　dbqq@phei.com.cn

通信地址：北京市万寿路 173 信箱

电子工业出版社总编办公室

邮　　编：100036

目　录

第 1 章　引　　论

学习要点

（1）数值计算的特点。计算方法研究的对象是用计算机求解各类数学问题，它既具有纯数学的抽象性与严密性特点，又具有应用的广泛性与实验的技术性特点。

（2）误差理论。误差的来源、误差的度量、误差的传播。

（3）数值计算的若干原则。避免两相近数相减和绝对值太小的除数，简化计算步骤，使用数值稳定的算法。

教学建议

本章是“计算方法”课程的开始，要求学生了解计算方法的特点和研究对象，重点学习误差的基本概念和性质，掌握绝对误差、相对误差和有效数字的关系，了解数值计算的基本原则。建议学时数为 2～4 学时。

1.1　从数学到计算

随着计算机的广泛应用和科学技术的高速发展，大量复杂的科学计算问题呈现在人们面前，要完成这些工作，仅靠人的自身努力是不可能的，必须借助于计算机这一人类有史以来最伟大的科技发明，而使计算机有效解决科学计算问题的关键技术是数值计算方法。让我们首先从下面的几个例子谈起。

引例 1（例 1.1.1）　设 $f(x)$ 在 $[a,b]$ 区间单调连续，并且 $f(a)\cdot f(b)<0$，求方程 $f(x)=0$ 在区间 $[a,b]$ 内的根。

解　由微积分的知识可知，方程 $f(x)=0$ 在 $[a,b]$ 区间内有唯一的实根，解此问题与函数 $f(x)$ 的形式有关。

若 $f(x)=ax^2+bx+c$，则由一元二次方程的求根公式知，方程 $f(x)=0$ 的根为

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

若 $f(x)=x^3-1=(x-1)(x^2+x+1)$，由代数知识可知，$f(x)=0$ 的根为

$$x_1=1,\ x_2=\frac{-1+\mathrm{i}\sqrt{3}}{2}=\omega,\ x_3=\frac{-1-\mathrm{i}\sqrt{3}}{2}=\omega^2 \qquad (\mathrm{i}^2=-1) \tag{1.1.1}$$

若 $f(x)=x^3+px+q$，可令 $x=\alpha+\beta$，代入得

$$(\alpha+\beta)^3+p(\alpha+\beta)+q=0$$

展开得

$$\alpha^3+3\alpha^2\beta+3\alpha\beta^2+\beta^3+p\alpha+p\beta+q=0$$

因式分解得 $(\alpha^3+\beta^3+q)+(\alpha+\beta)(3\alpha\beta+p)=0$，因为 $x=\alpha+\beta$ 不能为 0，可令

$$\alpha^3+\beta^3+q=0\text{，}\quad 3\alpha\beta+p=0$$

则 $\alpha\beta=-\dfrac{p}{3}$，记为 $\alpha^3\beta^3=-\dfrac{p^3}{27}$，以及 $\alpha^3+\beta^3=-q$。

由韦达定理，可将 α^3,β^3 看成是二次方程 $z^2+qz-\dfrac{p^3}{27}=0$ 的根，得

$$\alpha^3=\frac{-q}{2}+\sqrt{\frac{q^2}{4}+\frac{q^3}{27}}\text{，}\quad \beta^3=\frac{-q}{2}-\sqrt{\frac{q^2}{4}+\frac{q^3}{27}}$$

$$\alpha=\sqrt[3]{\frac{-q}{2}+\sqrt{\frac{q^2}{4}+\frac{q^3}{27}}}\text{，}\quad \beta=\sqrt[3]{\frac{-q}{2}-\sqrt{\frac{q^2}{4}+\frac{q^3}{27}}}$$

由此可得

$$x_1=\alpha+\beta=\sqrt[3]{\frac{-q}{2}+\sqrt{\frac{q^2}{4}+\frac{q^3}{27}}}+\sqrt[3]{\frac{-q}{2}-\sqrt{\frac{q^2}{4}+\frac{q^3}{27}}}\tag{1.1.2}$$

另外两个共轭复根为：$x_2=\alpha\omega+\beta\omega^2$，$x_3=\alpha\omega^2+\beta\omega$，其中 ω 为式（1.1.1）的值。

例如 $f(x)=x^3-x-1$，则由式（1.1.2），其实根为

$$x=\sqrt[3]{\frac{1}{2}+\sqrt{\frac{23}{108}}}+\sqrt[3]{\frac{1}{2}-\sqrt{\frac{23}{108}}}$$

若 $f(x)=ax^3+bx^2+cx+d$，则可将方程 $f(x)=0$ 两边除以 a，并设 $x=y-\dfrac{b}{3a}$，可将原方程化为如下形式

$$y^3+py+q=0$$

由此解出 y_1,y_2,y_3 后得 $x_i=y_i-\dfrac{b}{3a}$ $(i=1,2,3)$。

若 $f(x)$ 为四次多项式，可想而知，其求根公式会更加复杂，不过仍然可以求解。遗憾地是，挪威年轻的数学家阿贝尔（Abel）在 1826 年证明，若 $f(x)$ 为五次及以上的多项式，则代数方程 $f(x)=0$ 的根，不能由系数的初等函数表示，即五次及以上的代数方程无求根公式。

但是，在历史上，冥王星的发现归结为求解一个八次方程。不仅如此，人们还要求超越方程，例如确定轨道上行星位置的开普勒（Kepler）方程为

$$x=q\sin x+a\quad（其中 0<q<1，为椭圆轨道的偏心率）$$

实际上，用本书第 3 章介绍的牛顿法，可以求解此类方程。

引例 1 说明，某些数学问题在理论上没有解决的方法，解决此类问题必须用数值计算方法求其近似值。

引例 2（例 1.1.2） 求解 n 阶线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$。其中

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\text{，}\quad \boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}\text{，}\quad \boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

由线性代数的知识，根据克莱姆（Gramer）法则，若系数矩阵 $\boldsymbol{A}$ 是非奇异矩阵，即 $\det(\boldsymbol{A})\neq 0$，则线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一的解。

$$x_1 = \frac{\det(\boldsymbol{A}_1)}{\det(\boldsymbol{A})},\quad x_2 = \frac{\det(\boldsymbol{A}_2)}{\det(\boldsymbol{A})},\quad \cdots,\quad x_n = \frac{\det(\boldsymbol{A}_n)}{\det(\boldsymbol{A})} \tag{1.1.3}$$

其中矩阵

$$\boldsymbol{A}_j = \begin{pmatrix} a_{11} & \cdots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2(j-1)} & b_2 & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & b_n & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix} \quad (j = 1, 2, \cdots, n)$$

是用常向量b代换系数矩阵$\boldsymbol{A}$的第j列后所得的n阶方阵。式中，$a_{1(j-1)}$表示第1行第$j-1$列，其余类同。

这就意味着要解决此问题，需要计算$\det(\boldsymbol{A}), \det(\boldsymbol{A}_1), \det(\boldsymbol{A}_2), \cdots, \det(\boldsymbol{A}_n)$，共$n+1$个行列式，而每个行列式展开后有$n!$项，每项有$n$个数，需要$n-1$次乘法。如果忽略加、减、除法的次数，仅计算乘法的次数，则为$(n+1)\times n!\times(n-1)$。例如，当$n$=20时，乘法次数为$21\times 20!\times 19 \approx 9.7\times 10^{20}$。

若用每秒完成12.5万次乘除法的计算机计算，则需$9.7\times 10^{20} \div (1.25\times 10^5) \approx 7.8\times 10^{15}$秒。而1年$=365\times 24\times 3600 \approx 3.2\times 10^7$秒，若计算时间用年计算，则需要$7.8\times 10^{15} \div (3.2\times 10^7) \approx 2.4\times 10^8$年，约为两亿四千万年。

如果用每秒1亿次乘除法的银河I型巨型计算机计算此问题需30万年，用每秒10亿次乘除法的银河II型巨型计算机计算此问题需3万年，即使用目前世界上最快的“天河2号”计算机进行计算，也需要数月。而现代科学技术，如大型喷气客机、石油储藏模拟、流体涡度计算等，未知数常常超过百万。显然，仅仅只靠提高计算机的计算速度不是解决此类问题的主要手段。

实际上，用本书第4章介绍的高斯消去法，同样用每秒12.5万次乘除法的计算机求解20阶线性方程组只需要大约0.02秒的时间。

引例2说明，某些数学问题在理论上有解决的方法，但在实际中并不实用，必须寻找新的行之有效的计算方法。

引例3（例1.1.3） 求定积分$I_n = \int_0^1 \frac{x^n}{x+5}\mathrm{d}x \quad (n = 1, 2, \cdots, 20)$。

解 因为

$$\begin{aligned} I_n + 5I_{n-1} &= \int_0^1 \frac{x^n}{x+5}\mathrm{d}x + 5\int_0^1 \frac{x^{n-1}}{x+5}\mathrm{d}x = \int_0^1 \frac{x^n + 5x^{n-1}}{x+5}\mathrm{d}x = \int_0^1 \frac{x^{n-1}(x+5)}{x+5}\mathrm{d}x \\ &= \int_0^1 x^{n-1}\mathrm{d}x = \frac{1}{n}x^n\Big|_0^1 = \frac{1}{n} \end{aligned}$$

由此可得递推公式：

$$I_n = \frac{1}{n} - 5I_{n-1} \tag{1.1.4}$$

并且

$$I_0 = \int_0^1 \frac{1}{x+5}\mathrm{d}x = \ln(x+5)\Big|_0^1 = \ln 6 - \ln 5 = \ln\frac{6}{5} \approx 0.182322$$

由微积分的知识，我们知道I_n有如下性质：

① $I_n > 0$（因为被积函数$f(x)$在（0，1）上非负）

② I_n单调递减（当$n_1 > n_2$时，$I_{n_1} > I_{n_2}$）

③ $\lim\limits_{n\to\infty} I_n = 0$（因为$\lim\limits_{n\to\infty}\frac{x^n}{x+5} = 0$，$x\in[0,1]$）

④ $\frac{1}{6n} < I_{n-1} < \frac{1}{5n}$（$n > 1$）

由式（1.1.4）得：$I_{n-1}=\dfrac{1}{5n}-\dfrac{1}{5}I_n<\dfrac{1}{5n}$，又因为 I_n 单调递减 $I_n=\dfrac{1}{n}-5I_{n-1}<I_{n-1}$，故 $I_{n-1}>\dfrac{1}{6n}$。

现在用两种方法计算 I_n。

方法 A：按式（1.1.4）直接从 I_1 计算到 I_{20}，其计算结果见表 1.1.1。

表 1.1.1　计算结果 1

n	I_n	n	I_n	n	I_n	n	I_n
1	0.088 392 2	6	0.024 323 9	11	0.017 324 7	16	−10.1569
2	0.058 038 9	7	0.021 237 8	12	−0.003 290 22	17	50.8433
3	0.043 138 7	8	0.018 810 9	13	−0.093 374 2	18	−254.161
4	0.034 306 3	9	0.017 056 6	14	−0.395 442	19	1270.86
5	0.028 468 6	10	0.014 716 9	15	2.043 88	20	−6354.23

方法 B：由 I_n 的性质④知：$I_{n-1}\approx\left(\dfrac{1}{5n}+\dfrac{1}{6n}\right)/2$

所以
$$I_{20}\approx\frac{\dfrac{1}{6\times21}+\dfrac{1}{5\times21}}{2}=0.00873016$$

然后利用递推公式
$$I_{n-1}=\frac{1}{5n}-\frac{1}{5}I_n \tag{1.1.5}$$

自 I_{20} 计算到 I_1，其计算结果见表 1.1.2。

表 1.1.2　计算结果 2

n	I_n	n	I_n	n	I_n	n	I_n
19	0.008 253 97	14	0.011 229 2	9	0.016 926 5	4	0.034 306 3
18	0.008 875 52	13	0.012 039 9	8	0.018 836 9	3	0.043 138 7
17	0.009 336 01	12	0.012 976 6	7	0.021 232 6	2	0.058 038 9
16	0.009 897 50	11	0.014 071 3	6	0.024 325 0	1	0.088 392 2
15	0.010 520 5	10	0.015 367 6	5	0.028 468 4	0	0.182 322

现在我们对这两种方法进行比较，方法 A 的初值 I_0 具有 6 位有效数字，比较精确，但由该方法产生的数值解自 I_{12} 开始出现负值，且其绝对值逐渐增加，这显然与 I_n 的性质相矛盾，因此由方法 A 计算的结果不符合原问题的要求。而方法 B 的初始值 I_{20} 取的是一个近似的平均值，误差比较大，但所得结果不但完全符合原问题的要求，而且最后 I_0 的值与方法 A 的初始值相同，具有较高的精度。其根本的原因是方法 A 由式（1.1.4）直接计算，从 I_{n-1} 到 I_n 每向前推进一步，若 I_{n-1} 有误差，则其计算值的舍入误差增长 5 倍，误差的放大传播导致最终的结果与原问题的真值相悖，因此，是不稳定的。而方法 B 由式（1.1.5）从 I_n 到 I_{n-1} 计算，每向后推进一步，若 I_n 有误差的话，其舍入误差便减少为原来的 1/5，因此，获得了与原问题的性质一致的数值结果，是稳定的。

实际上，用本书第 1 章的后续知识，可以避免方法 A 现象的发生。

引例 3 说明，某些数学问题即使在实践中有解决的方法，仍然需要进行误差分析。

从以上 3 个引例可以看出，在解决科学计算问题时，经典的数学方法受到了极大的限制，虽然，近代数学家们提出了许多理论与方法，证明了一些问题的解存在、唯一以及解的某些特

征，但仍然只是对解的性质给出了某些定性的描述，而科技工作者和工程设计师却需要真实的、定量的数据。因此，许多数学问题的解决必须借助于计算机，而且要选择合理、有效的方法。我国著名的计算科学家石钟慈院士指出:“计算不仅仅只是作为验证理论模型的正确手段，大量的实例表明它已是重大科学发现的一种重要手段”，“科学计算与实验，理论三足鼎立，相辅相成，成为当今科学发现的三大方法”。

本书介绍的计算方法就是专门研究各种数学问题的计算机解法（数值解法），包括方法的构造和求解过程的理论分析及软件实现，它是计算数学的一个主要部分，包括方法的收敛性、稳定性以及误差分析等。计算方法既具有纯数学的抽象性与严密性的特点，又具有应用的广泛性与实验的技术性特点，因此，在学习计算方法时，要充分考虑计算机的特点，使所构造的算法，应该只包括计算机能直接处理的算术运算和逻辑运算，并严格控制计算的复杂性，最后还要在相关数学理论的基础上对误差进行分析。

由于数学的学科十分广泛，所出现的数学问题也各不相同。本书只涉及工程和科学实验中常见的数学问题，其中包括线性方程组、函数插值、微积分、微分方程、非线性方程、矩阵的特征值、函数逼近以及优化计算等，这些问题是解决其他数学问题的基础。

由于本书的内容包括了微积分、微分方程、非线性方程和线性方程组的数值计算方法，因此，在介绍这些常见的数值计算方法之前，第 2 章简要介绍了微积分、常微分方程和线性代数方面的基本内容，以便读者查阅。

1.2 误差理论初步

1.2.1 误差的来源

用数值计算方法求解数学问题，不可避免地会产生误差。实际上，在各种实际问题的求解过程中，误差的产生是绝对的，精确值却是相对的，产生误差的原因一般有以下几种。

1. 模型误差

例如求一个鸡蛋的表面积。首先要建一个数学模型，可以近似用球的表面积公式计算。但鸡蛋与球的形状差别较大，为了减少误差，可以用椭球的表面积公式计算。但鸡蛋的形状是一头大，一头小，仍然与椭球的形状不同，为了进一步减少误差，可以将鸡蛋的曲线画出来，利用曲面积分公式计算，但仍然会产生误差。这种**数学模型的解与实际问题的解之间出现的误差，称为模型误差。**

2. 测量误差

在建立数学模型以后，接下来就要进行一些数据的测量。例如，为了求鸡蛋的表面积，就要进行测量相关数据，若用球的表面积公式计算，则要测量其半径；若用椭球表面积计算，则要测量长半轴和短半轴；若要用曲面积分公式计算，则要测量积分区间等。由于测量手段的限制，因此在实际测量中，总会产生误差。这种**在测量具体数据时产生的误差称为测量误差。**

3. 截断误差（也称方法误差）

当用数学模型不能求出问题的精确解时，就要用数值计算方法求解，例如用梯形法求定积分

$$I = \int_a^b f(x)\,\mathrm{d}x$$

则公式为
$$I \approx I_1 = \frac{b-a}{2}[f(a)+f(b)] \tag{1.2.1}$$

实际上，这是用过两点 $(a, f(a)),(b, f(b))$ 的直线 $l(x)$ 代替被积函数 $f(x)$ 得到的积分值，即

$$I_1 = \int_a^b l(x)\,\mathrm{d}x$$

在本书第 8 章中，我们会知道 I 与 I_1 之间的误差为

$$R_1 = I - I_1 = -\frac{(b-a)^3}{12} f^{''}(\eta) \qquad \eta \in (a,b)$$

R_1 称为梯形求积公式的方法误差。这种**数学模型的准确解与数值计算方法的准确解之间的误差称为截断误差**。因为截断误差是方法固有的，所以又称为方法误差。对某数学问题的数值解进行误差估计，主要是指方法误差，这是本书要讨论的重点。

4．舍入误差

由于计算机字长的限制，某些数（如无理数 π）不能在计算机内精确表示，计算结果必然也会产生误差。例如，梯形求积公式（1.2.1）在用计算机求解 $f(a), f(b)$ 时，一般只能得到它们的近似值 $\tilde{f}(a)$ 和 $\tilde{f}(b)$，加上式（1.2.1）计算的误差，最终只能得到 I_1 的近似值 I_2。

$$I_1 \approx I_2 = \frac{b-a}{2}[\tilde{f}(a)+\tilde{f}(b)]$$

I_1 与 I_2 之间的误差就是舍入误差。这种**由于计算机字长的限制而产生的误差，称为舍入误差。**

针对不同的数值计算方法，误差估计的侧重点也不同。例如，在线性方程组的数值求解中，主要讨论输入数据的误差和舍入误差的传播，而在数值积分和微分中，重点分析各种方法的截断误差。

1.2.2 误差的度量

对于同一个数学问题，采用不同的方法会得出不同的结果。衡量某种方法优劣的标准之一，是看其结果的误差是否较小。一般度量误差的标准有三种形式。

（1）绝对误差与绝对误差限

定义 1.2.1 设 x 为某量的精确值，x^* 是它的一个近似值，则称 $E(x^*) = x - x^*$ 为近似值 x^* 的绝对误差，简称误差。由于精确值 x 是未知的，因此 x^* 的绝对误差 $E(x^*)$ 一般也是求不出来的。但是如果能求出 x^* 误差的一个范围 $E(x^*) = |x - x^*| \leqslant \delta(x^*)$，则称 $\delta(x^*)$ 为近似值 x^* 的绝对误差限，简称误差限。

例如，设 $x = \pi = 3.1415926535\cdots$，若取 x 的一个近似值 $x^* = 3.14159$，则

$$\delta(x^*) = |x - x^*| \leqslant 0.5\times 10^{-5}$$

称 x^* 的误差限为 0.5×10^{-5}。

一个近似数的误差限并不唯一，通常取满足 $|x - x^*| \leqslant \frac{1}{2}\times 10^n$（$n$ 为整数）的最小值。

（2）相对误差与相对误差限

绝对误差有时不能完全刻画一个近似数的精确程度，例如测量一个书桌和一个体育场的面积，误差都是 1cm^2，显然，后者的测量更精确，因此，决定某量的近似值的精度，除了考虑绝

对误差的大小外，还要考虑该量自身的大小，为此引入相对误差的概念。

定义 1.2.2 设 x 为某量的精确值，x^* 是它的一个近似值，则称 $E_r(x^*)=\dfrac{x-x^*}{x}$（$x\neq 0$）为 x^* 的相对误差。

由于精确值 x 是未知的，因此，在实际计算中常取 $E_r(x^*)=\dfrac{x-x^*}{x^*}$ 作为 x^* 的相对误差。若 $E_r(x^*)$ 的绝对值小于某个已知正数 $\delta_r(x^*)$，即 $\left|E_r(x^*)\right|=\left|\dfrac{x-x^*}{x^*}\right|\leqslant\delta_r(x^*)$，则称 $\delta_r(x^*)$ 为近似值 x^* 的相对误差限。

例如，设 $x=\pi=3.14159265358\cdots$，取 x 一个近似值 $x^*=3.14159$，则

$$\delta_r(x^*)=\left|\frac{x-x^*}{x^*}\right|\leqslant 0.8\times 10^{-6}$$

称 x^* 的相对误差限为 0.8×10^{-6}。

（3）有效数字

当某量的精确值的位数较多时，我们通常采用“四舍五入”的方法取 x 的前面若干位，作为 x 的近似值。

例如，$x=3.14159265358\cdots$

取 1 位	$x_1=3$	$\delta(x_1)\approx 0.14\leqslant 0.5$
取 5 位	$x_5=3.1416$	$\delta(x_5)\approx 0.000007<0.00005$
取 10 位	$x_{10}=3.141592654$	$\delta(x_{10})\approx 0.00000000042<0.0000000005$

这些近似值的误差限都不超过该近似值的最后 1 位数字的半个单位，则称它们都是有效数字，由此得有效数字的定义。

定义 1.2.3 如果近似值 x^* 的误差限不超过某位的半个单位，若该位数字到 x^* 的第一位非零数字共有 n 位，那么这 n 位数字称为 x^* 的有效数字，并称 x^* 具有 n 位有效数字。

在计算机中参加运算的数往往要进行规格化表示，因此，有效数字也可以定义如下。

定义 1.2.4 设 x^* 是 x 的一个近似值，写成规格化形式

$$x^*=\pm 10^k\times 0.a_1a_2\cdots a_n\cdots \tag{1.2.2}$$

式中 $a_i(i=1,2,\cdots)$ 是 0～9 之间的整数，且 $a_1\neq 0$，k 为整数。

如果

$$\left|x-x^*\right|\leqslant\frac{1}{2}\times 10^{k-n} \tag{1.2.3}$$

则称 x^* 为 x 的具有 n 位有效数字的近似值。

例 1.2.1 设 $x=\sqrt{200}=14.142\cdots$，$x^*=14.1$

$y=\lg 2=0.30102\cdots$，$y^*=0.3010$

$z=\mathrm{e}^{-5}=0.0067379\cdots$，$z^*=0.00673$

求各近似值的有效数字。

解 方法 1：

由于 $\delta(x^*)=\left|x-x^*\right|=0.042\leqslant 0.05$，小于十分位的半个单位，因此，14.1 每位都是有效数字，故 x^* 有 3 位有效数字。

由于 $\delta(y^*)=\left|y-y^*\right|=0.00002<0.00005$，小于万分位的半个单位，因此，0.3010 中小数点后